AD-A054 548   WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER   F/G 12/1
A VARIATIONAL EQUIVALENT TO DIAGONAL SCALING.(U)
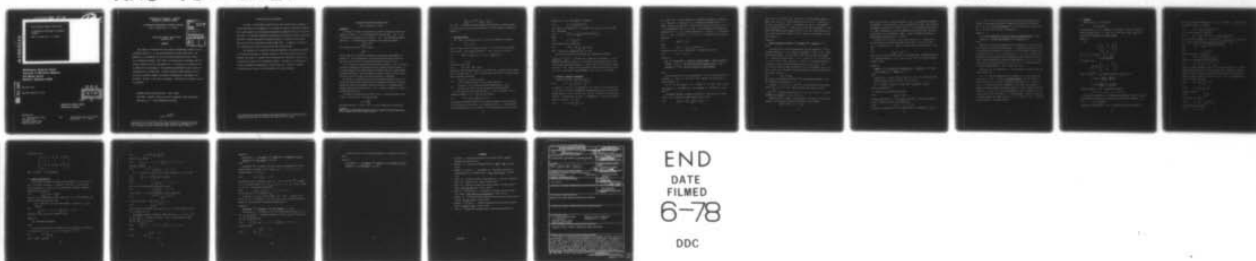FEB 78   M A BERGER, C T KELLEY   DAAG29-75-C-0024
UNCLASSIFIED   MRC-TSR-1829   NL

| OF |
AD
A054548

END
DATE
FILMED
6-78
DDC

MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

MRC Technical Summary Report #1829

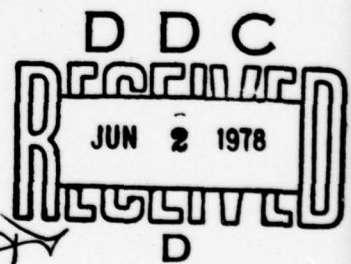# A VARIATIONAL EQUIVALENT TO DIAGONAL SCALING

Marc A. Berger and C. T. Kelley

Mathematics Research Center
University of Wisconsin—Madison
610 Walnut Street
Madison, Wisconsin 53706

February 1978

Received January 23, 1978

Approved for public release
Distribution unlimited

UNIVERSITY OF WISCONSIN - MADISON
MATHEMATICS RESEARCH CENTER

A VARIATIONAL EQUIVALENT·TO DIAGONAL SCALING

Marc A. Berger and C. T. Kelley

Technical Summary Report #1829
February 1978

ABSTRACT

This paper is concerned with the problem of diagonally scaling a given
nonnegative matrix a to one with prescribed row and column sums. The
approach is to represent one of the two scaling matrices as the solution
of a variational problem. This leads in a natural way to necessary and suf-
ficient conditions on the zero pattern of a so that such a scaling exists.
In addition the convergence of the successive prescribed row and column sum
normalizations is established. Certain invariants under diagonal scaling
are used to actually compute the desired scaled matrix, and examples are
provided. Finally, at the end of the paper, a discussion of infinite systems
is presented.

sec 1423

SIGNIFICANCE AND EXPLANATION

In order to avoid heavy round-off error when solving large systems of linear equations on the computer, it is useful to scale the coefficient matrix. This scaling amounts to multiplying the matrix by diagonal matrices on the left and right. We consider nonnegative coefficients and concern ourselves with the problem of choosing appropriate diagonal matrices in order that the scaled matrix have certain prescribed row and column sums. In addition, if some of the coefficients are zero, we ask if this can at all be done.

Our approach is to replace the scaling problem by an equivalent problem of finding the minimum of a certain function. This facilitates much of the analysis and leads in a natural way to necessary and sufficient conditions on the pattern of zeros in the coefficient matrix. Included is a technique for actually computing the appropriate scaling matrices, a number of examples illustrating the results, and a discussion of infinite systems.

A VARIATIONAL EQUIVALENT TO DIAGONAL SCALING

Marc A. Berger and C. T. Kelley

Introduction

In this paper we present a new approach to the problem of diagonally scaling a non-negative $m \times n$ matrix a to one with prescribed row and column sums σ and δ ; and to the question of convergence of the successive σ and δ row and column sum normalizations of a . Our methods show this equivalent to the problem of minimizing

$$\prod_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} t_i \right)^{\delta_j}$$

over $\mathbf{R}_+^m$ (positive vectors in $\mathbf{R}^m$) subject to

$$\prod_{i=1}^{m} t_i^{\sigma_i} = 1 .$$

And this leads us in a natural way to conditions on the pattern of zero elements in a (Theorem 2.D). The main computational result underlying the theoretical parts of the paper states that the solution to the above minimum problem is given by (3.12).

In §1 we present the problem and the tools employed to solve it. In §2 the existence, non-existence, uniqueness and convergence results are presented for finite systems. In §3 certain invariants under diagonal scaling are described, and used to actually compute the desired scaled matrix. And some examples are provided. Finally, in the last section a discussion of infinite systems, using functional analysis, is presented.

The results in §2 concerning the scaling and convergence questions are not new. The original work involved doubly stochastic matrices, and appears in Sinkhorn [7], Sinkhorn and Knopp [8], Menon [5], and Brualdi, Parter and Schneider [3]. The generalization to non-square matrices and arbitrary positive σ and δ appears in Bacharach [1], Brualdi [2], and Menon and Schneider [6].

The approach of representing the solution in terms of a variational problem is used by Theil [9] and Gorman [4]. The former minimizes

$$b_{ij} \log \frac{b_{ij}}{a_{ij}}$$

over positive matrices b having σ and δ as row and column sums. And the latter minimizes

$$\sum_{i=1}^{m} \sum_{j=1}^{n} a_{ij} e^{t_i + s_j} - \sum_{i=1}^{m} \sigma_i t_i - \sum_{j=1}^{m} \delta_j s_j$$

over $\mathbb{R}_+^{m+n}$. The problem we formulate, however, is tied directly to the Menon operator in [5], and leads in a most natural and direct way to the full admissibility condition of Theorem 2.D.

§1. The Scaling Problem

Let σ and δ be positive vectors in $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively, satisfying

$$(1.1) \qquad \sum_{i=1}^{m} \sigma_i = \sum_{j=1}^{n} \delta_j .$$

And let $\mathcal{Q}$ denote the set of nonnegative matrices a in $\mathbb{R}^m \otimes \mathbb{R}^n$ with no vanishing rows or columns. Define normalizing operators R and C from $\mathcal{Q}$ into itself by

$$(1.2) \qquad (Ra)_{ij} = a_{ij} (\sum_{\ell=1}^{n} a_{i\ell})^{-1} \sigma_i , \quad (Ca)_{ij} = a_{ij} (\sum_{k=1}^{m} a_{kj})^{-1} \delta_j .$$

Let

$$(1.3) \qquad \Lambda = CR .$$

The reader can verify that

$$(1.4) \qquad \mathcal{J}(\Lambda) = \mathcal{J}(C) \cap \mathcal{J}(R) \equiv \mathcal{J}(C,R)$$

over the set $\mathbb{R}_+^{m,n}$ of positive matrices in $\mathbb{R}^m \otimes \mathbb{R}^n$, where $\mathcal{J}$ denotes the set of fixed points. Of course, over $\mathcal{Q}$, $\mathcal{J}(\Lambda)$ is strictly larger than $\mathcal{J}(C,R)$. We concern ourselves with the following two problems.

(P_1) For which $a \in \mathcal{Q}$ does the sequence $\{\Lambda^N a\}$ converge to a limit in $\mathcal{J}(C,R)$, and to what limit?

(P_2) For which $a \in \mathcal{Q}$ do there exist positive diagonal matrices x and y such that $xay \in \mathcal{J}(C,R)$, and are they unique up to scalar multiples?

Let $\mathbb{R}_+^m$ denote the set of positive vectors t in $\mathbb{R}^m$. For a fixed $a \in \mathcal{Q}$ we associate an operator $\lambda \equiv \lambda(a)$ from $\mathbb{R}_+^m$ into itself by

$$(1.5) \qquad (\lambda t)_i = [\sum_{j=1}^{n} a_{ij} \delta_j (\sum_{k=1}^{m} a_{kj} t_k)^{-1}]^{-1} \sigma_i$$

This operator was defined by Menon in [5], and was used by Brualdi, Parter and Schneider in [3] to analyze (P_2) when $\mathcal{J}(C,R)$ is the assignment polytope Ω_n. It was also used by Menon and Schneider in [6] to analyze (P_2) in the more general setting put forth here.

-2-

We note that if $\lambda t = t$ then the matrix b defined by

$$(1.6) \qquad b_{ij} = a_{ij} \left(\sum_{k=1}^{m} a_{kj} t_k \right)^{-1} t_i \delta_j$$

is in $\mathfrak{J}(C,R)$. And conversely, if b of the form $b_{ij} = a_{ij} t_i s_j$ is in $\mathfrak{J}(C,R)$, then $\lambda t = t$. Furthermore,

$$(1.7) \qquad (\Lambda^N a)_{ij} = a_{ij} \left[\sum_{k=1}^{m} a_{kj} (\lambda^{N-1} t^0)_k \right]^{-1} (\lambda^{N-1} t^0)_i \delta_j$$

where

$$(1.8) \qquad t_i^0 = \left(\sum_{j=1}^{n} a_{ij} \right)^{-1} \sigma_i \ .$$

And by inverting this it follows that

$$(1.9) \qquad (\lambda^N t^0)_i = \sigma_i^{N+1} \prod_{k=0}^{N} \left[\sum_{j=1}^{n} (\Lambda^k a)_{ij} \right]^{-1}$$

Thus λ allows one to consider both (P_1) and (P_2).

We show in the next section that solving $\lambda t = t$ corresponds to minimizing

$\prod_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} t_i \right)^{\delta_j} \prod_{i=1}^{m} t_i^{-\sigma_i}$. This leads us in a natural way to conditions on σ and δ, relative to the pattern of zeros in a, such that (P_1) and (P_2) may be solved. In the third section we construct invariants of a under C and R which allow one to compute the limit of $\{\Lambda^N a\}$. Finally, in the last section we discuss the case where one or both of m and n are infinite, and solve analogs of (P_1) and (P_2) in this setting.

§2. Existence - Uniqueness - Convergence

Consider the function Φ defined on $\mathbb{R}_+^m$ by

$$(2.1) \qquad \Phi(t) = \prod_{j=1}^{n} \left(\sum_{i=1}^{m} a_{ij} t_i \right)^{\delta_j} \prod_{i=1}^{m} t_i^{-\sigma_i} \ .$$

The condition $\nabla \log \Phi(t) = 0$ is equivalent to $\lambda t = t$. Thus if Φ has a critical point in $\mathbb{R}_+^m$ then, in accord with the remarks in the previous section, there exist positive diagonal matrices x and y such that $xay \in \mathfrak{J}(C,R)$. We consider the problem of minimizing Φ. It follows from (1.1) that Φ is homogeneous in t, and thus it suffices to impose the constraint $t \in B$, where

$$(2.2) \qquad B = \{ t \in \mathbb{R}_+^m : \sum_{i=1}^{m} t_i = 1 \}$$

-3-

If $a \in \mathbb{R}_+^{m,n}$ then ϕ tends to infinity as t approaches the boundary ∂B from inside of B. And since ϕ is continuous in B, it follows that a minimum exists in B. However, if some of the elements of a vanish, then this need not be the case. In fact, if F is a proper subset of $\mathbb{Z}_m = \{1,\ldots,m\}$ and F' is its complement, we can examine the behavior of $\phi(t^* + \varepsilon\tau_\varepsilon)$ as ε tends to zero; where t_i^* is zero for $i \in F$ and positive for $i \in F'$, and $\tau_\varepsilon \in \mathbb{R}_+^m$ with $\tau_{\varepsilon,i} = 1$ for $i \in F$. The condition that the limit be infinite is

$$(2.3) \qquad \sum_{j \in C(F)} \delta_j < \sum_{i \in F} \sigma_i$$

where

$$(2.4) \qquad C(F) = \{ j : a_{kj} = 0 \text{ for each } k \in F' \} .$$

Furthermore, note that $C(F)$ is empty unless F contains one of the sets

$$(2.5) \qquad F_\ell = \{ i : a_{i\ell} > 0 \}$$

Theorem 2.A:

Let $a \in \mathcal{A}$ be such that (2.3) holds for all proper subsets F which are unions of the sets F_ℓ in (2.5). Then there exist positive diagonal matrices x and y such that $xay \in \mathcal{J}(C,R)$, and they are unique up to scalar multiples.

Proof:

First we establish (2.3) for any proper subset F. Let F^* be the union of the F_ℓ contained in F. Then $F^* \subset F$ and $C(F^*) = C(F)$. Therefore (2.3) must hold for F, since it holds for F^*.

Let $t^* \in \partial B$. Then $\lim_{\varepsilon \downarrow 0} \phi(t^* + \varepsilon\tau_\varepsilon) = \infty$ whenever $t^* + \tau_0 \in \mathbb{R}_+^m$. Thus it follows that $\lim_{t \to t^*} \phi(t) = \infty$. This means that ϕ has a minimum in B, and thus that x and y exist.

To establish uniqueness it suffices to show that if $\tilde{a}$ and $\tilde{x}\tilde{a}\tilde{y}$ are in $\mathcal{J}(C,R)$ where $\tilde{x}$ and $\tilde{y}$ are positive diagonal matrices then $\tilde{a} = \tilde{x}\tilde{a}\tilde{y}$.

$$(2.6) \qquad \sum_{i=1}^m \tilde{a}_{ij}(1 - \tilde{x}_i\tilde{y}_j) = \sum_{j=1}^n \tilde{a}_{ij}(1 - \tilde{x}_i\tilde{y}_j) = 0 .$$

-4-

Scale $\tilde{x}$ and $\tilde{y}$ so that each $\tilde{x}_i \le 1$ and each $\tilde{y}_j \ge 1$, and at least one element is equal to one. Say $\tilde{x}_{i_0} = 1$. Then it follows from (2.6) that $\tilde{y}_j = 1$ whenever $\tilde{a}_{i_0 j} > 0$. Thus $\tilde{a}_{i_0 j} = \tilde{a}_{i_0 j} \tilde{x}_{i_0} \tilde{y}_j$ for every j, and moreover, some $\tilde{y}_{j_0} = 1$. Then similarly $\tilde{a}_{i j_0} = \tilde{a}_{i j_0} \tilde{x}_i \tilde{y}_{j_0}$ for every i. That is, the i_0-th row and j_0-th column of $\tilde{a} - \tilde{x} \tilde{a} \tilde{y}$ vanish. So we delete them and consider the $(m-1) \times (n-1)$ case. Continuing this reduction, eventually either one or both of $\tilde{x}$ and $\tilde{y}$ reduce to a single element. And this case is trivial. ∎

Theorem 2.B:

<u>Under</u> <u>the</u> <u>conditions</u> <u>of</u> <u>Theorem</u> 2.A, <u>the</u> <u>sequence</u> $\{\Lambda^N a\}$ <u>converges</u> <u>to</u> $x\,a\,y$.

Proof:

It follows from (1.7) and the uniqueness result that the convergence of $\{\lambda^N t^0\}$ to an interior fixed point t of λ, implies the convergence of $\{\Lambda^N a\}$ to $x\,a\,y$. Suppose first that there exists p such that $(a\bar{a})^p \in \mathbb{R}_+^{m,n}$ where $\bar{a}$ denotes the transpose of a. Then for any distinct vectors u and v in $\mathbb{R}_+^m$ with $u \le v$, we not only have $\lambda u \le \lambda v$, but $(\lambda^p u)_i < (\lambda^p v)_i$ for every i. Now let t be an interior fixed point of λ, and let $[\alpha_0, \beta_0]$ be the smallest interval such that $\alpha_0 t \le t^0 \le \beta_0 t$. Then $\alpha_0 t \le \lambda t^0 \le \beta_0 t$ and we can reduce the interval $[\alpha_0, \beta_0]$ to the smallest $[\alpha_1, \beta_1]$ such that $\alpha_1 t \le \lambda t^0 \le \beta_1 t$. Continuing in this manner we generate a sequence of nested intervals $[\alpha_N, \beta_N]$ such that

(2.7) $$\alpha_N t \le \lambda^N t^0 \le \beta_N t$$

and $[\alpha_N, \beta_N]$ is the smallest such interval. This nested sequence converges to $[\alpha, \beta]$, and it suffices now to show that $\alpha = \beta$.

So suppose $\alpha \ne \beta$ and let $t^* = \lim_{k \to \infty} \lambda^{N_k} t^0$ be any cluster point of the sequence $\{\lambda^N t^0\}$. Then $\alpha t \le t^* \le \beta t$ and $[\alpha, \beta]$ is the smallest such interval. But then $\hat{t} = \lim_{k \to \infty} \lambda^{N_k + P} t^0$ is another cluster point, and $\alpha t_i < \hat{t}_i < \beta t_i$ for every i; and this contradicts the minimality of $[\alpha, \beta]$. Thus $\alpha = \beta$.

Suppose next that no powers of $a\bar{a}$ are positive. Then, since $a\bar{a}$ has a positive diagonal, it can be written as a direct sum of blocks b_q over subspaces

$$\mathbb{R}_q^m = \{t \in \mathbb{R}^m : t_i = 0, \ i \in F_q\}$$

-5-

where $F_q \subseteq \mathbb{Z}_m$, and each b_q has a positive power. And λ can accordingly be decomposed into operators λ_q acting on $\mathbb{R}^m_{q,+}$, the set of positive vectors in $\mathbb{R}^m_q$. Thus it suffices to establish the convergence of $\{\lambda^N_q t^{0,q}\}$, where $t^{0,q}$ is the corresponding restriction of t^0 to $\mathbb{R}^m_q$. Furthermore, this will follow immediately from the first part of this proof, as long as each λ_q has fixed points in $\mathbb{R}^m_{q,+}$. And this is immediate from Theorem 2.A, which establishes fixed points for λ in $\mathbb{R}^m$. $\blacksquare$

Theorems 2.A and 2.B solve (P_1) and (P_2) under conditions (2.3). In particular we note that under these conditions there always exists a matrix $b \in \mathfrak{F}(C,R)$ such that $a_{ij} = 0$ if and only if $b_{ij} = 0$. What we show below is that if there exists any F such that

(2.8)
$$\sum_{j \in C(F)} \delta_j > \sum_{i \in F} \sigma_i$$

then no such b exists, even if we allow b to have extra zeros. And therefore for these $a \in \mathcal{A}$ $\underline{no}$ diagonal matrices x and y exist for (P_2). And $\{\Lambda^N a\}$ does not converge to a limit in $\mathfrak{F}(C,R)$.

Theorem 2.C:

$\underline{Suppose}$ $a \in \mathcal{A}$ $\underline{is}$ $\underline{such}$ $\underline{that}$ (2.8) $\underline{holds}$ $\underline{for}$ $\underline{some}$ F . $\underline{Then}$ $\underline{for}$ $\underline{no}$ $b \in \mathcal{A} \cap \mathfrak{F}(C,R)$ $\underline{is}$ $\underline{it}$ $\underline{true}$ $\underline{that}$ $a_{ij} = 0$ $\underline{implies}$ $b_{ij} = 0$.

Proof:

Suppose such a b exists. Then (2.8) holds for b even if b has more zeros than a , because this will only serve to enlarge $C(F)$. But then

$$\sum_{i \in F} \sigma_i = \sum_{i \in F} \sum_{j=1}^{m} b_{ij} \geq \sum_{i \in F} \sum_{j \in C(F)} b_{ij} = \sum_{j \in C(F)} \sum_{i \in F} b_{ij} = \sum_{j \in C(F)} \sum_{i=1}^{m} b_{ij} = \sum_{j \in C(F)} \delta_j$$

which is a contradiction. $\blacksquare$

Finally we wish to consider the case where

(2.9)
$$\sum_{j \in C(F)} \delta_j = \sum_{i \in F} \sigma_i$$

for some proper subset(s) F , (2.3) holding for the others. In this case in order that x and y solving (P_2) exist, it is necessary that

(2.10)
$$a_{ij} = 0 , \quad i \in F , \quad j \in C(F)'$$

for those F satisfying (2.9). This is simply because it is necessary that $x_i a_{ij} y_j = 0$,

-6-

i ∈ F , j ∈ C(F)'. The result below shows that these conditions are sufficient as well. In order to avoid cumbersome hypotheses we summarize these conditions in a definition. Thus a ∈ 𝒶 will be called <u>fully</u> <u>admissible</u> if (2.3) or (2.9) hold for every proper sub-set F ; and (2.10) holds for those F satisfying (2.9).

Theorem 2.D:

a ∈ 𝒶 <u>is</u> <u>fully</u> <u>admissible</u> <u>if</u> <u>and</u> <u>only</u> <u>if</u> <u>there</u> <u>exist</u> <u>diagonal</u> <u>matrices</u> x <u>and</u> y <u>solving</u> (P_2). <u>And</u> <u>under</u> <u>these</u> <u>conditions</u> $\{\Lambda^N a\}$ <u>converges</u> <u>to</u> x a y .

Proof:

Sufficiency has already been established in the discussion above. To prove necessity we can assume (2.9) holds for some proper subset F . Otherwise the existence of x and y follows directly from Theorem 2.A. Because of (2.10) we can reduce the problem to the two sub-matrices $(a_{ij} : i \in F, j \in C(F))$ and $(a_{ij} : i \in F', j \in C(F)')$. Both of these sub-matrices are fully admissible and their dimensions are strictly less than those of a . Thus an induction establishes the existence of diagonal matrices x' and y', x" and y" solving (P_2) for these smaller matrices. And x',x" and y',y" can be pieced together to construct diagonal matrices x and y solving (P_2) for a .

The vector (x_i) *is a fixed point of* λ , and its existence is enough to justify the proof of Theorem 2.B. ∎

If a ∈ 𝒶 is not fully admissible, but is such that (2.3) or (2.9) hold for every proper subset F , then a is said to be <u>partially</u> <u>admissible</u>. For such a matrix a , there will not exist b ∈ 𝒶 ∩ 𝔍(C,R) having the same zero pattern as a . (But if we allow b to have extra zeros, then such matrices b will exist.) Thus Theorem 2.D amounts to saying that whenever there exists a matrix in 𝒶 ∩ 𝔍(C,R) having the same zero pattern as a , then x and y exist, and conversely. Hence φ defined by (2.1) has critical points in $\mathbb{R}_+^m$ if and only if a is fully admissible. Finally we note that for the doubly stochastic case, where m=n and $\sigma_i = \delta_j = 1$ for each i and j , full admissibility reduces to a cardinality condition.

-7-

§3. Invariants

Observe that if π is a k-cycle then

$$(3.1) \qquad \prod_{p=1}^{k} a_{i_p j_p} a_{i_p \pi_{j_p}}^{-1}$$

is invariant under C and R, provided it is well-defined. In fact it is invariant under any diagonal scaling $x\,a\,y$. These ratios can be used to compute $\lim_{N\to\infty} \Lambda^N a$ when it exists, as illustrated by the following example.

Example 3.A:

Let

$$a = \begin{pmatrix} 0 & a_{12} & a_{13} \\ a_{21} & 0 & a_{23} \\ a_{31} & a_{32} & 0 \end{pmatrix}$$

and $\sigma_i = \delta_j = 1$ for each i and j. Then

$$r = a_{12} a_{23} a_{31} a_{13}^{-1} a_{32}^{-1} a_{21}^{-1}$$

is invariant. Thus

$$\lim_{N \to \infty} \Lambda^N a = \begin{pmatrix} 0 & \mu & 1-\mu \\ 1-\mu & 0 & \mu \\ \mu & 1-\mu & 0 \end{pmatrix}$$

where $\mu = r^{\frac{1}{3}}\left(1+r^{\frac{1}{3}}\right)^{-1}$. And a diagonal scaling that leads to this is

$$x = \text{diag}\left(1 , \frac{a_{13}}{a_{23}} \frac{\mu}{1-\mu} , \frac{a_{12}}{a_{32}} \frac{1-\mu}{\mu}\right)$$

$$y = \text{diag}\left(\frac{a_{32}}{a_{31}a_{12}} \frac{\mu^2}{1-\mu} , \frac{\mu}{a_{12}}, \frac{1-\mu}{a_{13}}\right)$$

Finally, for general σ and δ conditions (2.3) become

$$\delta_1 < \sigma_2 + \sigma_3 , \quad \delta_2 < \sigma_1 + \sigma_3 , \quad \delta_3 < \sigma_1 + \sigma_2.$$

From now on we assume that a has a positive row and column. Without loss of generality these may be the first row and column. Then we can define a ratio matrix r by

$$(3.2) \qquad r_{ij} = a_{ij} a_{11} a_{i1}^{-1} a_{ij}^{-1} .$$

This matrix is invariant under C and R, and can be used to define analogues of λ

-8-

and ϕ. By writing the diagonal matrix x of (P_2) in the form $x_i = t_i a_{i1}^{-1}$ we are led to consider the operator $\hat{\lambda}$ defined by

$$(3.3) \qquad (\hat{\lambda}t)_i = [\sum_{j=1}^{n} r_{ij}\delta_j \, (\sum_{k=1}^{m} r_{kj}t_k)^{-1}]^{-1}\sigma_i.$$

If $\hat{\lambda}t=t$ then the matrix b defined by

$$(3.4) \qquad b_{ij} = r_{ij}(\sum_{k=1}^{m} r_{kj}t_k)^{-1}t_i\delta_j = a_{ij} \, (\sum_{k=1}^{m} a_{kj}t_k a_{k1}^{-1})t_i a_{i1}^{-1} \delta_j$$

is in $\mathcal{F}(C,R)$; and, conversely, if b of the form $b_{ij} = a_{ij}t_i a_{i1}^{-1}s_j$ is in $\mathcal{F}(C,R)$ then $\hat{\lambda}t=t$. And the corresponding relationship between iterates of $\hat{\lambda}$ and Λ is

$$(3.5) \qquad (\Lambda^N a)_{ij} = r_{ij}[\sum_{k=1}^{m} r_{kj}(\hat{\lambda}^{N-1}\hat{t}^0)_k]^{-1}(\hat{\lambda}^{N-1}\hat{t}^0)_i\delta_j$$

where

$$(3.6) \qquad \hat{t}_i^0 = a_{i1}t_i^0 \; .$$

And by inverting this it follows that

$$(3.7) \qquad (\hat{\lambda}^N\hat{t}^0)_i = a_{i1}\sigma_i^{N+1} \prod_{k=0}^{N} [\sum_{j=1}^{n} (\Lambda^k a)_{ij}]^{-1} \; .$$

Similarly the corresponding function to minimize is

$$(3.8) \qquad \hat{\phi}(t) = \prod_{j=1}^{n} (\sum_{i=1}^{m} r_{ij}t_i)^{\delta_j} \prod_{i=1}^{m} t_i^{-\delta_i} \; .$$

Problem (P_2) amounts to seeking a matrix b with the same ratio matrix as a, having prescribed row and column sums. And the uniqueness result in Theorem 2.A amounts to saying that distinct matrices with identical row and column sums have distinct ratio matrices. Finally, we note that because of (1.1) both ϕ and $\hat{\phi}$ are homogeneous in t, and thus it suffices to minimize

$$(3.9) \qquad \Psi(t) = \prod_{j=1}^{n} (\sum_{j=1}^{m} a_{ij}t_i)^{\delta_j}$$

or

$$(3.10) \qquad \hat{\Psi}(t) = \prod_{j=1}^{n} (\sum_{i=1}^{m} r_{ij}t_i)^{\delta_j}$$

over $\mathbb{R}_+^m$, subject to

$$(3.11) \qquad \prod_{i=1}^{m} t_i^{\sigma_i} = 1$$

The solution to the former is

$$(3.12) \qquad t_i^* = c \prod_{k=0}^{\infty} \sigma_i [\sum_{j=1}^{n} (\Lambda^k a)_{ij}]^{-1}$$

and the solution to the latter is

$$(3.13) \qquad \hat{t}_i^* = \hat{c} \, a_{i1}t_i^*$$

-9-

where c and ĉ are appropriate scale factors. The remainder of this section is devoted
to examples.

Example 3.B:

For the case   m=n=2   the limit is

$$\begin{pmatrix} \mu & \sigma_1-\mu \\ \delta_1-\mu & \mu+\sigma_2-\delta_1 \end{pmatrix}$$

where μ is a nonnegative solution of

$$\mu^2 - (\delta_1 + \frac{r_{22}\sigma_1 + \sigma_2}{r_{22}-1})\mu + \frac{r_{22}\sigma_1\delta_1}{r_{22}-1} = 0$$

provided a is nonsingular. Otherwise $\mu = \sigma_1\delta_1(\sigma_1+\sigma_2)^{-1}$. If $a_{22}=0$ the condition (2.3)
becomes $\delta_2 < \sigma_1$, and if $\delta_2 = \sigma_1$ then $\mu = 0$. The corresponding variational problem is to

$$\text{minimize} \quad (s+1)^{\delta_1}(s+r_{22})^{\delta_2}s^{-\sigma_1} .$$

Here . $s = t_1^{\frac{\sigma_1+\sigma_2}{\sigma_2}}$.

Example 3.C:

For the case   m=2, n=3   the limit is

$$\begin{pmatrix} \mu & \eta & \sigma_1-\mu-\eta \\ \delta_1-\mu & \delta_2-\eta & \mu+\eta+\delta_3-\sigma_1 \end{pmatrix}$$

where μ,η are nonnegative solutions of

$$\mu\eta + \frac{\delta_2}{r_{22}-1}\mu - \frac{r_{22}\delta_1}{r_{22}-1}\eta = 0$$

$$\mu^2 + \mu\eta - (\sigma_1 + \frac{r_{32}\delta_1+\delta_3}{r_{32}-1})\mu - \frac{r_{32}\delta_1}{r_{32}-1}\eta + \frac{r_{32}\sigma_1\delta_1}{r_{32}-1} = 0 .$$

These reduce to a single cubic in μ . If $r_{22} = 1$ then the first equation becomes
$\eta = \delta_2\delta_1^{-1}\mu$. If $r_{32} = 1$ then the second equation reduces to $(\delta_1 + \delta_3)\mu + \delta_3\eta = \sigma_1\delta_1$.
And if $r_{22} = r_{23} = 1$ then $\mu = \sigma_1\delta_1(\sigma_1 + \sigma_2)^{-1}$ and $\eta = \sigma_1\delta_2(\sigma_1+\sigma_2)^{-1}$. The corres-
ponding variational problem is to

$$\text{minimize} \quad (s+1)^{\delta_1}(s+r_{22})^{\delta_2}(s+r_{23})^{\delta_3}s^{-\sigma_1} .$$

-10-

Here, as before, $s = t_1^{\frac{\sigma_1+\sigma_2}{\sigma_2}}$.

Example 3.D:

For the case $m=n=3$, $\sigma_i = \delta_j = 1$ for each i and j ; if $r_{23}=r_{32}=r_{22}r_{33}$ or if $r_{22}=r_{33}=r_{23}r_{32}$ then the limit is a permuted matrix (i.e. all rows and columns are permutations of one another). In the former case it is symmetric; in the latter, skew-symmetric.

Example 3.E:

Let

$$a = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}$$

The conditions (2.3) become

$$\sigma_2 < \delta_1+\delta_4 \ , \quad \sigma_4 < \delta_1+\delta_2 \ , \quad \delta_2 < \sigma_1+\sigma_3 \ .$$

Thus we can consider the case $\sigma_i = \delta_j = 1$ for each i and j . The corresponding variational problem is to

minimize $(t_1+t_2+t_3+t_4)(t_1+t_3)(t_1+t_2+t_3)(t_1+t_3+t_4)$

subject to $t_1 t_2 t_3 t_4 = 1$.

By symmetry $t_1=t_3$ and $t_2=t_4$. Thus we need to minimize

$$(t_1^2 + 1)(2t_1 + t_1^{-1})^2 \ .$$

Letting $s = t_1^2$ the condition for a minimum is $8s^3 + 8s^2 = 1$. The only positive solution is $\frac{1}{4}(\sqrt{5}-1)$. Thus

$$t_1 = t_3 = \frac{1}{2}\sqrt{\sqrt{5}-1} \ , \ t_2 = t_4 = \sqrt{\sqrt{5}+1} \ .$$

This means that diagonal scaling matrices are

$$x = \text{diag}(t_1, t_2, t_3, t_4)$$

$$y = \frac{1}{2}\text{diag}(\sqrt{5\sqrt{5}-11}, \ t_2, \ \sqrt{2\sqrt{5}-4} \ , \ \sqrt{2\sqrt{5}-4} \ )$$

-11-

and the limit $x \, a \, y$ is

$$\begin{pmatrix} g - \frac{1}{2} & \frac{1}{2} & \frac{1}{2} - \frac{1}{2} g & \frac{1}{2} - \frac{1}{2} g \\[2mm] 1 - g & 0 & 0 & g \\[2mm] g - \frac{1}{2} & \frac{1}{2} & \frac{1}{2} - \frac{1}{2} g & \frac{1}{2} - \frac{1}{2} g \\[2mm] 1 - g & 0 & g & 0 \end{pmatrix}$$

where $g = \frac{1}{2} (\sqrt{5} - 1)$ is the golden mean.

§4. Scaling Infinite Matrices

In this section we consider the scaling of infinite matrices; i.e. one or both of m and n is infinite. We assume $n = \infty$, and that $\sigma \in \mathbb{R}^m$ and $\delta \in \ell_1$, are positive vectors satisfying (1.1). Denote by ℓ_p^m the Banach space of infinite $m \times \infty$-matrices a having real entries such that

(4.1)
$$\sum_{j=1}^{\infty} \sum_{i=1}^{m} |a_{ij}|^p \equiv \| a \|_{\ell_p^m}^p < \infty$$

Note that ℓ_2^m is a Hilbert space for any m (even $m = \infty$). Let $a \in \ell_1^m$ be positive, and consider the following analogue of (P_2):

Find $t \in \mathbb{R}_+^m$ and $s \in \mathbb{R}_+^\infty$ such that the matrix b given by $b_{ij} = a_{ij} t_i s_j$

satisfies

$(P_2)'$
$$\sum_{j=1}^{\infty} b_{ij} = \sigma_i \ , \quad i = 1, \dots, m \ ; \quad \sum_{i=1}^{m} b_{ij} = \delta_j \ , \quad j = 1, 2, \dots \quad .$$

The matrix $b \in \ell_1^m$ is said to be a solution of $(P_2)'$.

Theorem 4.A:

$(P_2)'$ has at most one solution.

Proof:

As in the proof of Theorem 2.A, we assume that $\tilde{a}$ and $(\tilde{t}_i \tilde{a}_{ij} \tilde{s}_j)$ are solutions of $(P)'$, and show that $\tilde{t}_i = t_1$; $i = 1, \dots, m$. For each $i = 1, \dots, m$

(4.2)
$$\tilde{t}_i = \sum_{k=1}^{m} \alpha_{ik} \tilde{t}_k$$

where $\alpha \in \mathbb{R}_+^{m,m}$ is given by

-12-

$$(4.3) \qquad \alpha_{ik} = \sigma_i^{-1} \tilde{t}_i \sum_{j=1}^{\infty} \tilde{a}_{ij} \tilde{a}_{kj} \delta_j^{-1} \tilde{s}_j \ .$$

We note that α satisfies

$$(4.4) \qquad \sum_{k=1}^{m} \alpha_{ik} = 1 \ , \quad i = 1,\ldots,m \ ; \quad \sum_{i=1}^{m} \alpha_{ik}\sigma_i = \sigma_k \ , \quad k = 1,\ldots,m \ .$$

By Jensen's inequality

$$(4.5) \qquad \gamma_i \equiv e^{-\tilde{t}_i} \leq \sum_{k=1}^{m} \alpha_{ik}\gamma_k \ ; \quad i = 1,\ldots,m \ .$$

If $m = \infty$, note that the sum on the left of (4.5) converges, as $\gamma_i \leq 1$. We have

$$(4.6) \qquad \sum_{i=1}^{m} \sigma_i \gamma_i \ \leq \ \sum_{k=1}^{m} (\sum_{i=1}^{m} \alpha_{ik}\sigma_i) \gamma_k = \sum_{k=1}^{m} \sigma_k \gamma_k \ ,$$

and so

$$(4.7) \qquad \gamma_i = \sum_{k=1}^{m} \alpha_{ik}\gamma_k \ ; \quad i = 1,\ldots,m \ .$$

But of course the same argument can be used to show that

$$(4.8) \qquad \gamma_i^n = \sum_{k=1}^{m} \alpha_{ik}\gamma_k^n \ ; \quad i = 1,\ldots,m$$

for any integer $n \geq 0$. And by taking limits of sums it follows that

$$(4.9) \qquad e^{-(\gamma_i - \gamma_1)^2} = \sum_{k=1}^{m} \alpha_{ik} e^{-(\gamma_k - \gamma_1)^2} \ .$$

If $\tilde{t}_i \neq \tilde{t}_1$ for some i, then for all $i = 1,\ldots,m$

$$(4.10) \qquad \sum_{k=1}^{m} \alpha_{ik} e^{-(\gamma_k - \gamma_1)^2} < 1 \ .$$

Since, by (4.9), the inequality (4.10) cannot hold for $i=1$, it must indeed be true that $\tilde{t}$ is a constant vector. $\blacksquare$

To consider the question of existence, suppose first that $m < \infty$. For $N \geq 1$ let $\rho_N = \sum_{i=1}^{m} \sigma_i \Big/ \sum_{j=1}^{N} \delta_j$. According to Theorem 2.A there is a unique matrix $a^N \in \mathbb{R}_+^{m,N}$ such that $a_{ij}^N = a_{ij} t_i^N s_j^N$ and

$$(4.11) \qquad \sum_{j=1}^{N} a_{ij}^N = \sigma_i \ , \quad i = 1,\ldots,m \ ; \quad \sum_{i=1}^{m} a_{ij}^N = \rho_N \delta_j \ , \quad j = 1,2,\ldots \ .$$

Define

$$(4.12) \qquad b_{ij}^N = \begin{cases} a_{ij}^N \ , & j \leq N \\[2ex] 0 \ , & j > N \end{cases}$$

Theorem 4.B:

For the case $m < \infty$, the sequence $\{b^N\}$ defined by (4.12) converges in the norm topology of ℓ_1^m to the solution b of $(P_2)'$.

Proof:

The sequence $\{b^N\}$ is bounded in ℓ_2^m and as such has a subsequence $\{b^{N_k}\}$ which converges weakly to an element $b \in \ell_2^m$. Clearly $b_{ij} \geq 0$; $i = 1,\ldots,m$, $j = 1,2,\ldots$. Define functionals f_j on ℓ_2^m by

$$(4.13) \qquad f_j(z) = \sum_{i=1}^{m} z_{ij}, \quad z \in \ell_2^m .$$

By taking weak limits it follows that $f_j(b) = \delta_j$. Thus $b \in \ell_1^m$ and $\{b^{N_k}\}$ converges to b in the norm topology of ℓ_1^m . And $\sum_{j=1}^{\infty} b_{ij} = \sigma_i$; $i = 1,\ldots,m$. As in §3 define r_{ij} by (3.2). Then, again by taking limits,

$$(4.14) \qquad b_{ij}b_{11} = r_{ij}\, b_{i1}\, b_{1j}; \quad i = 1,\ldots,m, \quad j = 1,2,\ldots .$$

And this shows that b is a solution to $(P_2)'$. ∎

If we write $b_{ij} = a_{ij} t_i s_j$ then, as before, $\lambda t = t$, where λ is given by (1.5). And if R, C and Λ are defined on ℓ_1^m by (1.2) and (1.3), then the same proof used for Theorem 2.B can be used here to establish the following result.

Theorem 4.C:

The sequence $\{\Lambda^N a\}$ converges in ℓ_1^m to the solution, b, of $(P_2)'$.

For the case $m = \infty$ the solution to $(P_2)'$ can be constructed as follows. For $n \geq 1$ set $\kappa_N = \sum_{j=1}^{\infty} \delta_j \Big/ \sum_{i=1}^{N} \sigma_i$. According to Theorem 4.A there is a unique $a^N \in \ell_1^N$ of the form $a_{ij}^N = a_{ij} t_i^N s_j^N$ such that

$$(4.15) \qquad \sum_{j=1}^{\infty} a_{ij}^N = \kappa_N \sigma_i, \quad i = 1,\ldots,N; \qquad \sum_{i=1}^{N} a_{ij}^N = \delta_j, \quad j = 1,2,\ldots .$$

As before, define

$$(4.16) \qquad b_{ij}^N = \begin{cases} a_{ij}^N, & i \leq N \\[2mm] 0, & i > N \end{cases}$$

-14-

An argument similar to the one used to prove Theorem 4.B establishes the following result.

Theorem 4.D:

For the case $m = \infty$, the sequence $\{b^N\}$ defined by (4.16) converges in the norm topology of ℓ_1^∞ to the solution b of $(P_2)'$.

-15-

REFERENCES

1. Bacharach, M., "Biproportional Matrices and Input-Output Change." Cambridge University Press, New York, 1970.

2. Brualdi, R. A., Convex sets of nonnegative matrices. Canad. J. Math. 20 (1968), 144-157.

3. Braauldi, R. A., Parter, S. V., and Schneider, H., The diagonal equivalence of a nonnegative matrix to a stochastic matrix. J. Math. Anal. and Appl. 16 (1966), 31-50.

4. Gorman, W. M., "Estimating trends in Leontief matrices" : A note on Mr. Bacharach's paper. Note. Nuffield College, Oxford, Duplicated, 1963.

5. Menon, M. V., Reduction of a matrix with positive elements to a doubly stochastic matrix. Proc. Amer. Math. Soc. 18 (1967), 244-247.

6. Menon, M. V., and Schneider, H., The spectrum of a nonlinear operator associated with a matrix. Linear Algebra and its Applications 2 (1969), 321-334.

7. Sinkhorn, R., A relationship between arbitrary positive matrices and doubly stochastic matrices. Ann. Math. Stat. 35 (1964), 876-879.

8. Sinkhorn, R., and Knopp, P., Concerning nonnegative matrices and doubly stochastic matrices. Pacific J. Math. 21 (1967), 343-348.

9. Theil, H., "Economics and Information Theory." North-Holland, Amsterdam, 1967.

MAB&CTK/db

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER (14) 1829 √ MRC-TSR-1829	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER (9) Technical

4. TITLE (and Subtitle)

A VARIATIONAL EQUIVALENT TO DIAGONAL SCALING.

5. TYPE OF REPORT & PERIOD COVERED

Summary Report, no specific reporting period

6. PERFORMING ORG. REPORT NUMBER

7. AUTHOR(s)

(10) Marc A. Berger and C. T. Kelley

8. CONTRACT OR GRANT NUMBER(s)

(15) DAAG29-75-C-0024
NSF-MCS75-17385

9. PERFORMING ORGANIZATION NAME AND ADDRESS

Mathematics Research Center, University of
610 Walnut Street   Wisconsin
Madison, Wisconsin 53706

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS

2 - Other Mathematical Methods

11. CONTROLLING OFFICE NAME AND ADDRESS

See Item 18

12. REPORT DATE

(11) February 1978

13. NUMBER OF PAGES

16   (12) 20 p.

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)

15. SECURITY CLASS. (of this report)

UNCLASSIFIED

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

U. S. Army Research Office   National Science Foundation
P.O. Box 12211   Washington, D. C. 20550
Research Triangle Park
North Carolina 27709

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Diagonal scaling, Jensen's inequality, Weak convergence.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

    This paper is concerned with the problem of diagonally scaling a given non-negative matrix a to one with prescribed row and column sums. The approach is to represent one of the two scaling matrices as the solution of a variational problem. This leads in a natural way to necessary and sufficient conditions on the zero pattern of a so that such a scaling exists. In addition the convergence of the successive prescribed row and column sum normalizations is established. Certain invariants under diagonal scaling are used to actually compute the desired scaled matrix, and examples are provided. Finally, at the end of the paper, a discussion of infinite systems is presented.

DD FORM 1473 1 JAN 73   EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

221 200

ED

78